frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 19 March 2012
doi: 10.3389/fpsyg.2012.00074

Gesture facilitates the syntactic analysis of speech

Henning Holle[1]*, Christian Obermeier[2], Maren Schmidt-Kassow[3], Angela D. Friederici[4], Jamie Ward[5] and Thomas C. Gunter[4]*

[1] *Department of Psychology, University of Hull, Hull, UK*
[2] *Minerva Research Group "Neurocognition of Rhythm in Communication", Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany*
[3] *Institute of Medical Psychology, Johann Wolfgang Goethe University, Frankfurt am Main, Germany*
[4] *Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany*
[5] *School of Psychology and Sackler Centre for Consciousness Science, University of Sussex, Brighton, UK*

Edited by:
Gabriella Vigliocco, University College London, UK

Reviewed by:
Sotaro Kita, Birmingham University, UK
Jeremy I. Skipper, Hamilton College, USA

****Correspondence:***
Henning Holle, Department of Psychology, University of Hull, Cottingham Road, HU6 7RX Hull, UK.
e-mail: h.holle@hull.ac.uk;
Thomas C. Gunter, Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 13, 04103 Leipzig, Germany.
e-mail: gunter@cbs.mpg.de

Recent research suggests that the brain routinely binds together information from gesture and speech. However, most of this research focused on the integration of representational gestures with the semantic content of speech. Much less is known about how other aspects of gesture, such as emphasis, influence the interpretation of the syntactic relations in a spoken message. Here, we investigated whether beat gestures alter which syntactic structure is assigned to ambiguous spoken German sentences. The P600 component of the Event Related Brain Potential indicated that the more complex syntactic structure is easier to process when the speaker emphasizes the subject of a sentence with a beat. Thus, a simple flick of the hand can change our interpretation of who has been doing what to whom in a spoken sentence. We conclude that gestures and speech are integrated systems. Unlike previous studies, which have shown that the brain effortlessly integrates semantic information from gesture and speech, our study is the first to demonstrate that this integration also occurs for syntactic information. Moreover, the effect appears to be gesture-specific and was not found for other stimuli that draw attention to certain parts of speech, including prosodic emphasis, or a moving visual stimulus with the same trajectory as the gesture. This suggests that only visual emphasis produced with a communicative intention in mind (that is, beat gestures) influences language comprehension, but not a simple visual movement lacking such an intention.

Keywords: language, syntax, audiovisual, P600, ambiguity

INTRODUCTION

When we talk to one another, communication does not only take place in the auditory domain, but also simultaneously in the visual domain. Conversational gestures, which are movements of the hands that co-occur with speech, are one important example of such cross-modal communication. The fact that conversational gestures are reliably elicited during spontaneous speech (even when talking on the phone, Bavelas et al., 2008) suggests that gestures serve an important communicative function that cannot be completely achieved by speech alone. So why do we gesture when we speak? Research addressing this question has either been looking for functions that gesture may have in the speaker (where they may facilitate the act of speaking, Krauss, 1998) or in the listener (where gesture may convey additional information not found in speech). With respect to the latter, it has been found that – contrary to some initial negative findings (Krauss et al., 1991, 1995) – listeners are sensitive to the additional information provided by gesture. For instance, several groups have found that semantically incongruent gesture–speech pairings interfere with language comprehension, and we have reported evidence that gestures can disambiguate lexically ambiguous words (Holle and Gunter, 2007; Obermeier et al., 2011). Thus, the semantic information provided by gestures interacts with the semantic information of speech, and recent brain imaging studies have implicated that the left inferior frontal gyrus (Willems et al., 2007, 2009) and the left posterior temporal lobe (Holle et al., 2008, 2010) are crucially involved in this interaction.

However, these insights relate mainly to representational gestures (e.g., gesturing *writing* with a pencil while saying *write*). In comparison, the relationship between speech and non-representational gestures (of which beat gestures are an important sub-category) is much less well understood. Beats are short, rhythmic hand movements that match the cadence of speech (McNeill, 1992; Hubbard et al., 2009). These gestures accompany speech pervasively and appear not to be under intentional control (Alibali et al., 2001). With respect to their potential communicative function, it has been suggested that they accent or emphasize portions of their co-expressive speech (Efron, 1941/1972). But how does such gestural emphasis influence our interpretation of a speaker's utterance? And if gestural emphasis does have an effect on language comprehension, would this effect then be specific to gesture, or would other forms of emphasis (e.g., pitch accents, visual movement) have the same impact? We hypothesized that beats might help a listener to figure out who is doing what to whom in sentences that are temporarily ambiguous with respect to their syntactic structure.

THE PRESENT STUDY

Our paradigm exploited the quite flexible word order of German, which allows expression of the same meaning (e.g., "The dog bites the mailman") in either a subject-initial word order ("[Der Hund]$_{Subj}$ beisst [den Postboten]$_{Obj}$") or an object-initial word order ("[Den Postboten]$_{Obj}$ beisst [der Hund]$_{Subj}$"). Based on this phenomenon of word order variation, we created sentences that were temporarily ambiguous with respect to their subject and object role. Consider the example sentence provided in **Figure 1**. Up to the sentence-final word, there are two possible interpretations (at least in German). First, it could be that the woman has greeted the men (assumed argument order: Subject–Object–Verb, SOV). Alternatively, it could also be that the men have greeted the woman (argument order: OSV). Only at the sentence-final word it becomes clear who has actually been doing what to whom. Note that SOV and OSV structures are not treated as equally probable in German[1]. Instead, there is a strong preference to analyze an ambiguous initial noun phrase as the subject of the sentence (Haupt et al., 2008). Therefore, a disambiguation toward the OSV structure is somewhat unexpected and elicits additional processing costs which can be observed at the disambiguating word (a) on a behavioral level as increased reading times (Schriefers et al., 1995; Bader and Meng, 1999) and (b) on an electrophysiological level as an increased P600 (Osterhout and Holcomb, 1992; Friederici et al., 1993; Hagoort et al., 1993; Knoeferle et al., 2008). The P600 is a positive-going deflection of the event related brain potential (ERP) peaking around 600 ms after the onset of the critical word. It is reliably elicited whenever an ambiguous input is disambiguated toward the syntactically more complex alternative (for review, see Haupt et al., 2008). Although there are several suggestions about the functional significance of the P600, one scheme these proposals have in common is that of reanalysis, be it terms of a specific *syntactic* reanalysis (Friederici, 2002) or a more general reanalysis, including perceptual errors (van de Meerendonk et al., 2010).

The remainder of the manuscript describes three EEG experiments that investigate in how far different types of emphasis cues (Experiment 1: visual beats; Experiment 2; the auditory pitch accents normally associated with visual beats; Experiment 3: the visual movement associated with beats) can help to reduce or even abolish the P600 effect usually associated with a disambiguation toward the more complex OSV structure. Such a reduction would imply that emphasis cues may bias toward (or prevent a de-selection of) an alternative syntactic structure.

EXPERIMENT 1: SEEING BEATS

Beat gestures are inherently multimodal. In addition to the obvious visual component, making a beat also affects the speech that it accompanies (increased pitch, increased duration, see Krahmer and Swerts, 2007). In Experiment 1, we were specifically interested in the impact of the visual component on syntactic disambiguation, while controlling for auditory differences. Therefore, we assembled spoken sentences taken from a non-beating speaker that were phonologically identical until the sentence-final word (see below) and combined them with a video of a speaker either producing no gesture, a beat emphasizing the first noun phrase (NP1), or a beat emphasizing the second noun phrase (NP2). In this way, we ensured that all observed ERP effects can only be due the presence of a visual beat, and not to their associated auditory pitch accents.

METHODS

Participants

Thirty-five German-speaking students participated in Experiment 1 after giving written informed consent following the guidelines of the Ethics committee of the University of Leipzig. Two participants had to be excluded due to excessive EEG artifacts, and nine because they had an overall error rate on the behavioral task exceeding 40%. The remaining 24 participants (12 female, mean 25 years of age, range 21–29) were right-handed (mean laterality coefficient 95.7, Oldfield, 1971). All participants had normal or corrected-to-normal vision and none reported any known hearing deficit.

Stimuli

All sentences consisted of a matrix clause, followed by a complement clause in perfect tense. The noun in the matrix clause was always a proper name. All verbs in the complement clause were transitive verbs requiring a direct accusative object. Of the two noun phrases in the complement clause, one NP was always a feminine singular, whereas the other NP was always plural (masculine, feminine, or neuter), making the two NPs case-ambiguous (either nominative or accusative case). This case ambiguity made all sentences temporarily ambiguous with respect to their syntactic structure. Sentences were only disambiguated at the sentence-final auxiliary verb, either toward a preferred SOV or a non-preferred OSV structure (for a stimulus example, see **Figure 1**).

The experimental sentences were created out of a set of 240 noun phrase–noun phrase–verb combinations (NP–NP–V). All verb participles had the form of *ge* + *verb stem*, which is a very common form of verb participle generation in German. No participle was repeated within the set of NP–NP–V combinations.

Recording and splicing. First, a recording list was assembled, containing only the subject-initial versions of the 240 NP–NP–V sets in two different variations: (1) A subject-initial structure, where the singular NP was followed by the plural NP (e.g., *Peter knows, that the woman the men greeted has*) and (2) a reversed subject-initial structure, where the plural NP was followed by the singular NP (e.g., *Peter knows, that the men the woman greeted have*). This was possible because the NPs of all sentences were carefully selected to contain as little semantic bias as possible, which allowed us to create semantically plausible reversed versions of each sentence.

Next, a professional speaker produced all stimuli from the recording list with a natural prosody without producing any beat gestures. We instructed the speaker to produce the stimuli with a broad focus, i.e., she was told to not highlight a particular word in the sentence. Each stimulus was produced at least twice. The acoustic signal of each sentence was visually inspected and acoustically tested to determine the time of six events: (1) Onset of the

[1]The standard word order in English is SVO, as in "I like that" and English does not allow the flexible word order variation discussed here where object and subject switch places. It is, however, possible to move the object to the front of the sentence in English, as in "*That* I like" (OSV). Such topicalizations can be used to place special emphasis on the object.

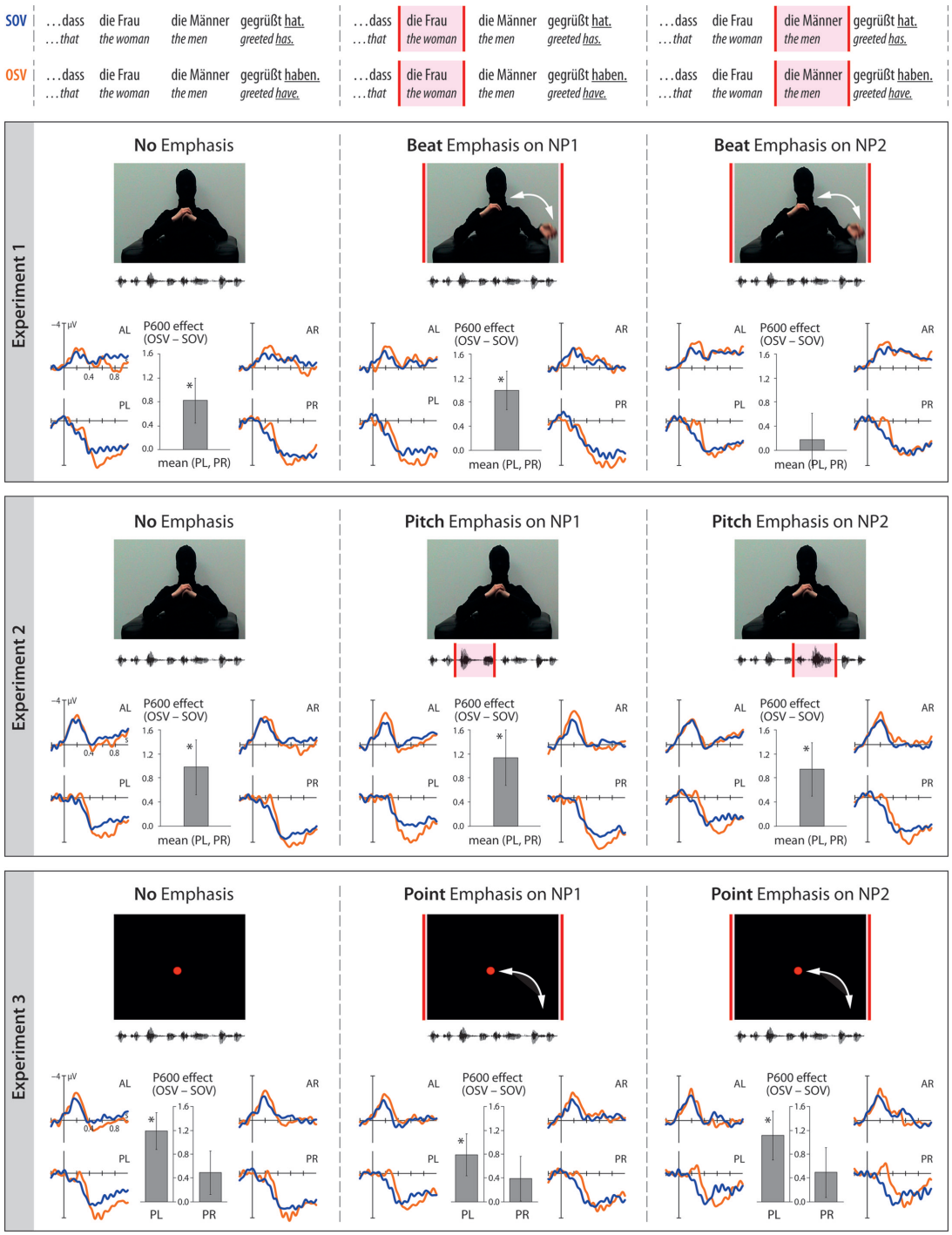


FIGURE 1 | Materials and Results. Original German sentences, as uttered in our video stimuli, as well as literal English translation in *italics*. Full German SOV sentence: Peter sagt, dass die Frau die Männer gegrüßt hat. English gloss: Peter says that the woman has greeted the men. Full German OSV sentence: Peter sagt, dass die Frau die Männer gegrüßt haben. English gloss: Peter says that the men have greeted the woman. ERPs were time-locked to the critical sentence-final words (underlined). ERPs for the preferred Subject–Object–Verb order (SOV) are shown in blue and the ERPs for the less preferred Object–Subject–Verb order (OSV) are shown in yellow. Grand-average ERPs (Experiment 1: $n = 24$; Experiment 2: $n = 19$; Experiment 3: $n = 23$) were averaged across four regions of interest: Anterior-Left (AL), Anterior-Right (AR), Posterior-Left (PL), and Posterior-Right (PR). Text highlighted by red bars indicates those portions of speech emphasized either by a visual beat gesture (Experiment 1), a beat-induced pitch accent (Experiment 2) or a moving point (Experiment 3). Bar graphs show the amplitude of the P600 effect (±SEM). ROIs in which the P600 effect is significant at $p < 0.05$ are marked by an asterisk (*).

determiner of NP1 (2) primary stress of NP1 (3) onset of the determiner of NP2 (4) primary stress of NP2 (5) onset of the participle and (6) onset of the sentence-final auxiliary.

Using this timing information, the audio recordings described in the previous paragraph were then recombined in a cross-splicing procedure detailed in **Figure 2**. In short, we extracted segments of the recorded sentences on the basis of the available timing information, and recombined these segments. The outcome of the splicing procedure were four different structures of each experimental item: (1) an SOV structure, where the first NP was singular and the second NP plural (SOV: $NP1_{sg}$–$NP2_{pl}$), (2) an SOV structure, where the order of the NPs was reversed (SOV: $NP1_{pl}$–$NP2_{sg}$), (3) an OSV structure, in which the first NP was singular and the second NP plural (OSV: $NP1_{sg}$–$NP2_{pl}$), and (4) an OSV structure with a reversed order of the two NPs (OSV: $NP1_{pl}$–$NP2_{sg}$, see **Figure 2**). Finally, all cross-spliced sentences were normalized to the same average sound pressure level using PRAAT (Boersma and Weenink, 2010). These stimuli formed the auditory component of our audiovisual stimuli for Experiment 1.

Gesture-recording and audiovisual synchronization. We re-invited the actress to record the beats gestures used for Experiment 1. She was placed in a comfortable chair with armrests and uttered a subset of the experimental sentences, either with no beat gesture, a left hand beat gesture accentuating the first NP, or a left hand beat gesture accentuating the second NP of the sentence. We also recorded versions where the right hand accentuated either the first or the second NP. Our gestures always started and ended in the central gesture space. They consisted of a rapid lowering of the forearm, a quick wrist movement at the apex, and a return to the resting to the resting position in the central space. Thus, although some degree of iconicity is present (according McNeill, 1992; e.g., more than two movement phases), they have many key characteristics of a beat gesture (beginning and end in central gesture space, wrist movement at the apex of the gesture). To minimize the influence of facial cues, the face of the actress was covered with a nylon stocking.

The video recordings were then combined with the previously created cross-spliced audio recordings. In the beat videos, the apex of the beat movement was always synchronized with the primary stress of the respective NP, resulting in a natural synchronization of gesture and speech (Levelt et al., 1985; McNeill, 1992). Each video contained 600 ms of silence before the onset of the sentence, where the speaker was also not moving. Example videos are provided as Supplementary Material.

As has been mentioned previously, the experimental stimuli were based on a set of 240 NP–NP–V combinations. Each item existed in 20 different versions, which differed with respect to their Structure (SOV or OSV), Number_Order (singular–plural or plural–singular), Emphasis (no beat, beat on NP1, beat on NP2), and Side (left hand beat, right hand beat). Thus, the total stimulus set for Experiment 1 consisted of 4800 video clips. These video clips were divided into 24 experimental lists, ensuring that each participant saw only one version of an item. Thus, all item-specific effects were counterbalanced across participants.

PROCEDURE

The participants were seated in a dimly lit, sound-attenuated chamber facing a computer screen. The videos were centered on a black background and extended for 10° visual angle horizontally and 8° vertically. A trial started with a fixation cross on the screen, which was presented for 300 ms, followed by the video presentation. Two seconds after sentence offset, a question concerning the content of the preceding sentence was presented until the participant responded with a button press ("yes" vs. "no") or until 3 s had elapsed. This question always had the form "*Was NP1/NP2 verbed?, e.g., Was the woman greeted?*". Each response was immediately followed by a feedback stimulus (500 ms) that informed the subjects about the accuracy of the response (correct/incorrect). The next trial began 1700 ms later with the presentation of a fixation cross.

An experimental session (excluding time for electrode application) lasted approximately 60 min. Stimuli were presented in four blocks each consisting of 60 items. Key assignment for

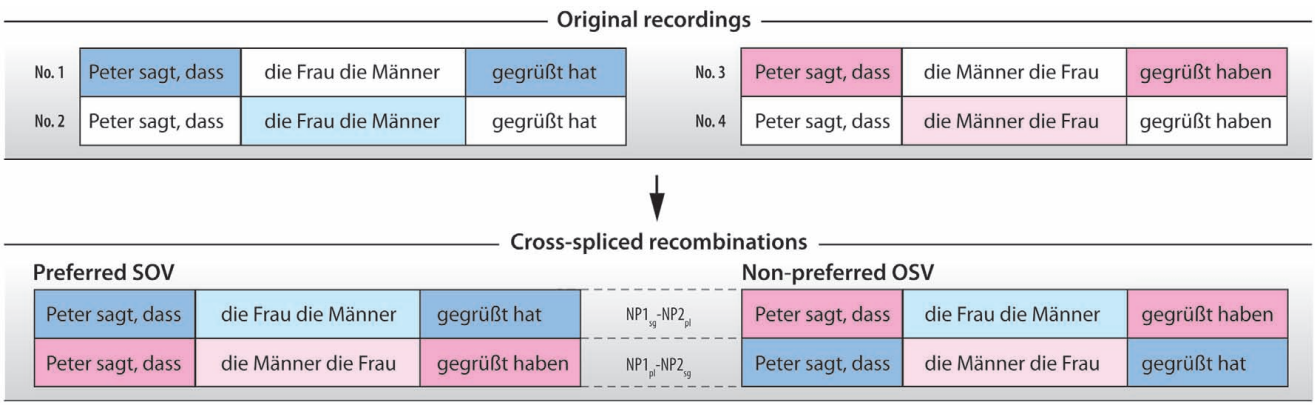


FIGURE 2 | Outline of the cross-splicing procedure for Experiment 1. Literal translation of the sentence on the left (portions used for cross-splicing are highlighted in blue): *Peter says, that the woman the men greeted has.* Literal translation of sentence on the right (highlighted in orange): *Peter says, that the men the woman greeted have.*

correct (left or right) was counterbalanced across participants. Each participant received one of the 24 experimental lists.

ERP recording

The EEG was recorded from 63 Ag/AgCl electrodes (Electro-Cap International). It was amplified using a PORTI-32/MREFA amplifier (DC to 135 Hz) and digitized online at 500 Hz. Electrode impedance was kept below 5 kΩ. Data were re-referenced offline to linked mastoids. Vertical and horizontal electrooculograms (EOG) were also measured.

Data analysis

Single-subject ERPs were calculated for each of the six experimental conditions. The epochs were time-locked to the onset of the disambiguating sentence-final auxiliary verb and lasted from 200 ms pre-stimulus onset to 1000 ms post-stimulus onset. A 200-ms pre-stimulus baseline was used. Four regions of interest (ROIs) were defined: anterior-left (AL): AF7, AF3, F7, F5, F3, FT7, FC5, FC3; anterior-right (AR): AF4, AF8, F4, F6, F8, FC4, FC6, FT8; posterior-left (PL): TP7, CP5, CP3, P7, P5, P3, PO7, PO3; posterior-right (PR): CP4, CP6, TP8, P4, P6, P8, PO4, PO8. An automatic artifact rejection using a 200-ms sliding window was performed on the EOG channels (±30 μV) and on the EEG channels (±40 μV) and was double-checked by visual inspection. Overall, approximately 20% of the trials did not enter statistical analysis due to artifacts or incorrect responses. Based on visual inspection of the data, a time window from 200 to 350 ms was used to analyze the early negativity, whereas a time window of 500–800 ms was used for the analysis of the P600 effects.

The statistical analysis of the ERP data was performed in two steps. The first analysis targeted the processing strategy in the absence of any visual cue. This analysis was implemented by means of a repeated-measures ANOVA using the within-subject factors Structure (SOV, OSV), Region (anterior, posterior), and Hemisphere (left, right). The second analysis explored how the brain response to the experimental sentences is modulated by a concurrent emphasis cue. This was done by means of a repeated-measures ANOVA with the factors Structure (SOV, OSV), Emphasis (NP1, NP2), Region (anterior, posterior), and Hemisphere (left, right).

Only effects that involve the critical factors of Structure or Emphasis are reported. Before entering statistical analysis, the data were filtered offline with a high-pass filter of 0.2 Hz. For presentation purposes only, an additional 10-Hz low pass filter was used.

RESULTS

Behavioral data

Generally, participants were very accurate in responding to the target questions (87% correct). The first analysis that targeted the behavior in the absence of a gesture cue indicated that accuracy was lower for OSV (83.1%) than for SOV (92.3%) structures, as indicated by a significant main effect of Structure [$F(1, 23) = 19.5$, $p < 0.0001$]. The second analysis tested how the presence of a beat gesture modulated the behavioral performance. The corresponding ANOVA with the factors Structure (SOV, OSV), Beat (NP1, NP2) revealed only a main effect of Structure [$F(1, 23) = 29.56$, $p < 0.0001$], which was due to a lower accuracy for object-initial (83.8%) as compared to subject-initial structures (91.5%). The main effect of Beat as well the interaction of Structure and Beat were not significant (both Fs < 1).

ERP data

Our first analysis of the ERP data targeted again the processing mode in the absence of a particular emphasis cue. The top left part of **Figure 1** shows the ERPs time-locked to the onset of the disambiguating auxiliary verb when the speaker did not produce an accompanying beat gesture. In this case, processing an object-initial as compared to a subject-initial structure elicited an early negativity in the time window from 200 to 350 ms, followed by a late positivity in the time window from 500 to 800 ms. Whereas the early negativity is broadly distributed across the scalp, the late positivity appears to be maximal at left posterior electrodes (see also Supplementary Material). On the basis of its scalp distribution, polarity, and latency, the late positivity was identified as a P600.

The statistical analysis for the early time window (200–350 ms) revealed that the early negativity was more pronounced for OSV than for SOV structures [$F(1, 23) = 5.16$, $p < 0.05$]. For the P600 time window (500–800 ms), the corresponding ANOVA yielded a significant main effect of Structure [$F(1, 23) = 4.64$, $p < 0.05$] indicating that the P600 was more pronounced when sentences were disambiguated toward the non-preferred object-initial structure.

To sum up, processing OSV as compared to SOV structures in the absence of a gesture cue elicited a broadly distributed early negativity, followed by a broadly distributed P600.

Next, we looked how these ERP patterns are modulated by the presence of a beat gesture. When the speaker produced a beat gesture on the first ambiguous noun phrase of the sentence, the ERP pattern appears to be similar to the pattern observed without a gesture: An early, broadly distributed negativity, followed by a P600 (see **Figure 1**, top center). However, when the beat accentuated the second ambiguous noun phrase, only an early negativity appears to be present, whereas the P600 effect is virtually absent (see **Figure 1**, top right).

The ANOVA for the early negativity yielded significant main effects of Structure [$F(1, 23) = 20.71$, $p < 0.0001$] and Beat [$F(1, 23) = 6.60$, $p < 0.05$], but no interaction between these two factors ($F < 1$). The main effect of Structure indicated that the negativity was more pronounced for object-initial as compared to subject-initial sentences, whereas the main effect of Beat reflected a generally more negative ERP when the beat fell on the second as compared to the first ambiguous NP.

The statistical analysis for the P600 time window revealed a significant main effect of Beat [$F(1, 23) = 12.85$, $p < 0.005$], a significant two-way interaction of Structure by Region [$F(1, 23) = 4.58$, $p < 0.05$] as well as a significant three-way interaction of Structure by Region by Hemisphere [$F(1, 23) = 5.01$, $p < 0.05$]. Most importantly, there was also a three-way interaction of Structure by Beat by Region [$F(1, 23) = 5.71$, $p < 0.05$]. *Post hoc* tests indicated that the P600 effect for OSV as compared to SOV structures was only significant at posterior sites when the beat accentuated the first NP [$F(1, 23) = 9.57$, $p < 0.01$], but not when it accompanied the second NP [$F(1, 23) < 1$].

To test whether a beat on the second ambiguous noun phrase might interfere with the processing of the SOV structure, we directly compared the P600 elicited by SOV structures presented without a beat with the P600 for SOV structures accompanied by a beat on the second noun phrase. No significant main effect of Beat or interactions involving this factor were observed (all $F < 2.99$, all $p > 0.1$), indicating that a beat on the second noun phrase did not interfere with the processing of the standard SOV order.

Finally, we tested whether the lack of a P600 effect when beats emphasized NP2 might be partly driven by an increase in the preceding negativity. Although the interaction of Beat by Structure was not significant for the preceding negativity (see above), **Figure 1** suggests that at least numerically, the early negativity is more pronounced for beats on NP2 than for beats on NP1. To directly test this possibility, we compared the ERP difference of (NP1: OSV–SOV) with (NP2: OSV–SOV) in the early time window at posterior sites (where the P600 modulation was significant). The corresponding paired t-test was not significant [$t(23) = 1.18$, $p = 0.25$], suggesting that there are no reliable difference in the early negativity between beat on NP1 and beat on NP2. Thus, what happens during the time window of the early negativity cannot explain what happens later on during the P600 time window.

One question the results from Experiment 1 raise is why we observed an interaction of Beat by Structure in the ERPs, but not in the behavioral data. In the behavioral results we found a strong main effect of Word Order, which was due to a lower accuracy score for OSV as compared to SOV structures. This may be seen as further evidence that additional processing costs arise when temporarily ambiguous syntactic structures are disambiguated toward the non-preferred object-initial structure (Schriefers et al., 1995; Bader and Meng, 1999). Beyond that, however, we believe that the behavioral data should be interpreted with caution. To avoid a contamination of the ERP data through the motor preparation and execution of the response, the question prompting participants to respond was displayed 2 s after the offset of the video. This leaves a considerable period of time for slower, offline processes (e.g., metalinguistic reasoning) to kick in, which most likely have influenced the response of our participants. The ERPs are in this sense a purer measure, because they provide a direct reflection of the online processes taking place at the disambiguating region of the sentence.

To summarize the results from Experiment 1, only the P600, but not the early negativity was modulated by the presence of a beat gesture. When either no beat accompanied the sentence or the beat fell on the first NP, we observed strong P600 effects. However, when the beat highlighted the second NP, the P600 effect was abolished.

EXPERIMENT 2: HEARING BEATS

In the first experiment, we observed that visual beat gestures can abolish the P600 usually associated with syntactically more complex sentences. Given that beat gestures tend to influence the speech they accompany (increased pitch and duration, see Krahmer and Swerts, 2007), one obvious question is now whether these auditory pitch accents induced by a beat gesture do also abolish the P600 effect. Therefore, we designed another experiment, which in some respect is a mirror version of Experiment 1. We took the speech of a speaker either not producing a beat gesture, or producing a beat on either one of two noun phrases, and paired the speech with a video of a non-gesturing speaker. Thus, visual stimulation is perfectly controlled for, and all modulations of the standard structure effect can only be attributed to the auditory manipulation.

METHODS

Participants

Twenty German-speaking students participated in Experiment 2. The data of one participant had to be excluded based on rejection criteria (i.e., excessive artifacts or an overall error rate on the behavioral task exceeding 40%). The remaining 19 participants (10 female, mean 25 years of age, range 19–28) were right-handed (mean laterality coefficient 90.6, Oldfield, 1971). All participants had normal or corrected-to-normal vision and none reported any known hearing deficit.

Stimuli

Because we were also interested in the extent to which auditory pitch accents associated with beat gestures also influence syntactic disambiguation, we included a condition in the recording session where the speaker was asked to make a beat gesture synchronized either with NP1 or NP2 while uttering the sentences. To create the materials for Experiment 2, we cross-spliced the NPs that were emphasized by a beat gesture into sentences not containing an emphasis cue, thereby creating a local manipulation of auditory emphasis induced by a beat gesture. Otherwise, the logic of the cross-splicing procedure was the same as outlined above (see **Figure 3** for details).

For Experiment 2, we wanted to make sure that in all experimental items the experimental manipulation was clearly audible. Therefore, the first author listened to all cross-spliced stimuli, and only those stimuli were included in the final set, in which all 12 auditory versions of each item [Structure (SOV, OSV) × Number_Order (sgl.–pl or pl–sgl.) × Emphasis (no emphasis, NP1 emphasis, NP2 emphasis)] could be correctly classified into one of the three emphasis conditions. He was blind with respect to the respective emphasis condition whilst listening to the stimuli. The final set for Experiment 2 consisted of 40 NP–NP–V combinations. Since each item consisted of 12 different auditory versions (see above), the total stimulus set for Experiment 2 consisted of 480 video clips. These stimuli were divided into two experimental lists, where each participant saw each item six times, once in each of the six possible combinations of the factors Structure (2) and Emphasis (3). The factor Number_Order was counterbalanced across the two lists.

The stimuli in Experiment 2 were presented in six Blocks, each consisting of 40 items. All other experimental details were as in Experiment 1.

RESULTS

Behavioral data

Again, participants had a lower accuracy for OSV than for SOV structures, both when processed in the absence of an emphasis cue [87.8 vs. 92.2%; $F(1, 18) = 8.02$, $p < 0.05$] as well as when processed in the context of a beat-induced pitch accent [87.8 vs.

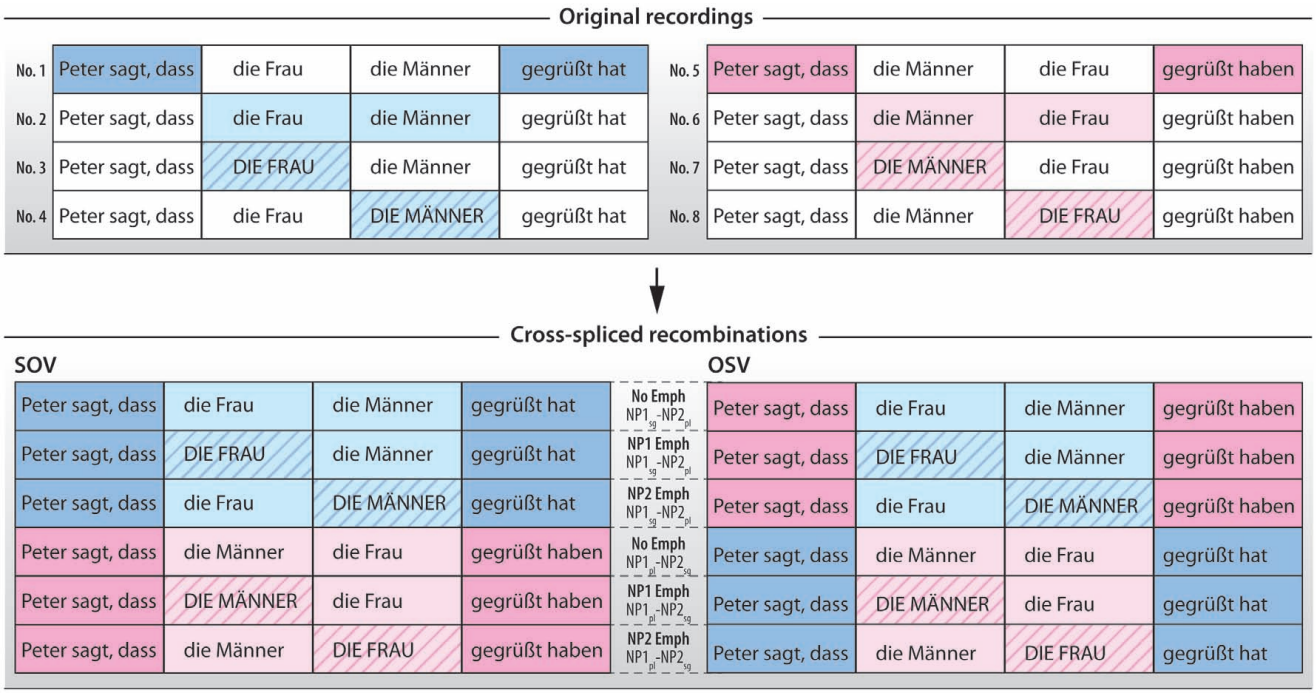


FIGURE 3 | Outline of the cross-splicing procedure for Experiment 2. Noun phrases in CAPITALS indicate those phrases during which the speaker produced a beat gesture during the recording.

94.7%; $F(1, 18) = 12.19$, $p < 0.005$]. No Structure by Emphasis interaction was observed [$F(1, 18) < 1$].

ERP data

The ERPs showed the familiar bi-phasic pattern. A disambiguation toward the less preferred OSV structures elicited an increased negativity, followed by a P600 effect. These effects appear not to be modulated by the presence or absence of auditory emphasis.

In the statistical analysis for the sentences without an emphasis cue, we observed no statistically significant effects in the early time window [but note that there was a trend toward a main effect of Structure ($F(1, 18) = 3.01$, $p = 0.09$)]. In the P600 time window, we observed a significant main effect of Structure [$F(1, 18) = 6.84$, $p < 0.05$].

Next we looked at how these effects are modulated by the presence of a beat-induced pitch accent emphasizing either NP1 or NP2. For the early time window, we obtained only a significant main effect of Structure [$F(1, 18) = 6.82$, $p < 0.05$], but no significant Structure by Emphasis interaction [$F(1, 18) < 1$]. Similarly, the analysis for the P600 time window revealed a statistical trend toward a main effect of Structure [$F(1, 18) = 3.98$, $p = 0.06$], a significant main effect of Emphasis [with Emphasis on NP1 more positive than on NP2; $F(1, 18) = 4.7$, $p < 0.05$], but no interaction between Structure and Emphasis [$F(1, 18) = 1.81$, $p = 0.19$].

To sum up, beat-induced pitch accents do not seem to interact with syntactic processing. We observed no significant interactions between Emphasis and Structure in any of our dependent variables (accuracy, early negativity amplitude, P600 amplitude).

EXPERIMENT 3: SEEING MOVEMENT

In Experiment 1, we have seen that visual beats on the second noun phrase can abolish the P600 effect usually associated with the more complex OSV structures. In Experiment 2, it was found that the auditory correlate of a beat does not produce this effect. Thus, the facilitation of OSV structures seems to depend on visual, rather than auditory emphasis. But is this really a gesture-specific effect, or is the facilitation simply due to the fact of seeing visual movement that is synchronized with speech? In a final experiment, we looked at this issue, by presenting a dot that either remained stationary in the center of the screen, or that followed the trajectory of the gesturing hand. These stimuli were paired with the (prosodically uninformative) speech used in Experiment 1.

METHODS

Participants

Twenty-four German-speaking students participated in Experiment 3 after giving written informed consent following the guidelines of the Ethics committee of the University of Leipzig. Data from one participant had to be excluded based on rejection criteria (i.e., excessive artifacts or an overall error rate on the behavioral task exceeding 40%). The remaining 23 participants (12 female, mean 25 years of age, range 20–30) were right-handed (mean laterality coefficient 94.4, Oldfield, 1971). All participants had normal or corrected-to-normal vision and none reported any known hearing deficit.

Stimuli

Experiment 3 was conducted to explore in how far visual movement synchronized with speech interacts with syntax. Therefore,

we measured the position of the gesturing hand in each video using a previously established method (Holle et al., 2008). The position of the relevant hand was recorded as the pixel coordinate of the junction point between index finger and thumb (or estimated, if occluded from sight), separately for each video frame. Next, we created a video in which a red dot moved along the trajectory defined by the pixel coordinates. This video, which showed either a stationary (in the case of video without a gesture) or a moving dot (in the case of a video that contained a beat gesture) was then combined with the prosodically uninformative speech stimuli that were created for Experiment 1. The audio–visual synchronization was identical to Experiment 1 (see above). We based Experiment 3 on the reduced item set of items used for Experiment 2; however, an additional 8 items were used to allow an even counterbalancing of stimulus effects across lists. Thus, Experiment 3 consisted of 48 NP–NP–V combinations and there were 20 different versions of each item (see also Experiment 1), resulting in a total of 960 video clips. These stimuli were divided into four experimental lists, where each participant saw each item six times, once in each of the six possible combinations of the factors Structure (2) and Emphasis (3). The factors Number_Order and Side were counterbalanced across the four lists.

The stimuli in Experiment 3 were presented in 6 Blocks, each consisting of 48 items. All other experimental details were as in Experiment 1.

RESULTS

Behavioral data

As before, participants' task performance was more accurate after SOV than after OSV sentences, independent of whether these sentences had been accompanied by an emphasizing point movement [93.2 vs. 84.2%; $F(1, 22) = 21.54$, $p < 0.0001$] or not [92.0 vs. 84.2%; $F(1, 22) = 11.09$, $p < 0.005$]. There was no Structure by Emphasis interaction [$F(1, 22) < 1$].

ERP data

The ERPs obtained in Experiment 3 show once again the familiar pattern. The more difficult OSV structures elicit an increased negativity, followed by an increased P600. The P600 effect seems to be a bit reduced when the moving points emphasize the first noun phrase.

In the absence of an emphasis cue, we observed a main effect of Structure in the time window for the early negativity [$F(1, 22) = 16.4$, $p < 0.0003$]. In the P600 time window, we observed a significant Structure by Hemisphere interaction [$F(1, 22) = 11.96$, $p < 0.005$], indicating that the P600 effect was more pronounced in the left hemisphere.

Next, we looked at in how far these ERP effects are modulated by the presence of a moving point emphasizing either the first or the second noun phrase. In the corresponding analysis for the early time window, we observed a main effect of Structure [$F(1, 22) = 21.25$, $p < 0.0001$], which additionally interacted with Region [$F(1, 22) = 9.63$, $p < 0.01$]. In the P600 time window, we obtained a significant two-way interaction of Emphasis by Region [$F(1, 22) = 5.23$, $p < 0.05$] as well as a significant three-way interaction of Structure by Region by Hemisphere [$F(1, 22) = 4.89$, $p < 0.05$]. No interactions involving both critical factors of Structure and Emphasis were observed [all $F(1, 22) < 1.27$, all $p > 0.27$]. Nonetheless, **Figure 1** suggests that the P600 effect is at least numerically larger at posterior sites in condition NP2. We explicitly tested this possibility by comparing the P600 amplitude difference of (OSV–SOV) at posterior sites between conditions NP1 and NP2. The corresponding paired t-test was not significant [$t(22) = 0.75$, $p = 0.45$] suggesting that there is no reliable difference in the amplitude of the P600 effect.

GENERAL DISCUSSION

We provide experimental evidence that gestural emphasis, by means of a beat gesture, influences which syntactic structure is assigned to a spoken sentence. In particular, we observed that the P600 effect, reflecting processing cost for non-preferred syntactic structures, is abolished when a beat gesture emphasizes the noun phrase that later turns out to be subject of the sentence. The effect was not observed for the auditory pitch accents associated with beats or a visual control condition. This novel finding has implications for recent theorizing about gesture and speech as an integrated system (Kelly et al., 2010b), multimodal theories of language comprehension (Crocker et al., 2010) and the role of visual attention in language comprehension.

As has been mentioned in the introduction, a P600 effect reflects reanalysis costs triggered by a disambiguation toward the syntactically more complex alternative. Our results are in line with this reanalysis view of the P600. Whenever participants' realize at the sentence-final word that their initial assumption of a subject-initial word order has been wrong, additional processing costs arise (reflected in a P600 effect) because the initial analysis has to be revised. The only exception to this rule is when the second ambiguous noun phrase of the sentence is accompanied by a beat gesture. The additional emphasis provided by the beat gesture seems to increase the plausibility of the OSV structure, so that it is just as available as the standard SOV structure (see Results).

Why is it specifically a beat on the second noun phrase that increases the plausibility of OSV, but not a beat on the first noun phrase? Possible explanations are related either to the timing between beat and the disambiguating element or to the syntactic view of NP1 most likely being the subject. The timing explanation is based on two assumptions. The first is that beats indicate newsworthy information (see also McNeill, 1992). In our particular experiment it was newsworthy when the sentences had the less frequent OSV structure. The second assumption of the timing explanation is that the impact of a beat is short-lived. A beat on NP1 may initially increase the plausibility of OSV, but this emphasis might have been decayed by the time the disambiguating sentence-final word is encountered. In contrast, in the case of a beat on NP2, there is less intervening time between the gestural emphasis and the disambiguating word, and therefore less decay. The syntax-related explanation is that the beats are treated as a cue to the subject of a sentence. According to this view, a beat on NP1 does not eliminate the P600, because it only provides the redundant information that the first NP1 should be interpreted as the subject of the sentence. This is because OSV structures when standing in isolation are quite infrequent in German (Kempen and Harbusch, 2003) and there is a strong tendency

to analyze an initial ambiguous argument as the subject of the sentence (Haupt et al., 2008). In contrast, when a beat emphasizes NP2, this provides the non-redundant information that the subject is in an unusual sentence position. Further experiments using a paradigm where an object is expected in the initial sentence position are needed to decide which of these two interpretations is correct.

In Experiment 2, we observed that pitch accents associated with beat gestures do not increase the plausibility of OSV structures as the visual beats do. This dissociation between visual beats and auditory accents may reflect that visual beats are a clearer and less variable emphasis cue than pitch accents. An isolated visual beat in a sentence is always a clear-cut emphasis cue. The listener will try to interpret this cue and come up with a plausible interpretation of why the speaker considers the phrase that accompanies the beat as newsworthy. In comparison, a pitch accent is a much less salient emphasis cue, since there is never just one isolated pitch accent within a sentence, but a multitude of major and minor accents. Furthermore, Haupt et al. (2008) have found that different speakers use very different prosodic patterns to mark object-initial patterns, and these speaker-specific variations make pitch accents a less valid cue for structural disambiguation during sentence comprehension.

In Experiment 3, we found that the facilitation of OSV structures is not just simply a visual attention effect. Moving dots that follow the exact trajectory of the beat gestures and are synchronized in the same way with the spoken sentences do not increase the plausibility of the object-initial sentences. This is surprising, because the moving dots are most likely more attention-capturing than the beat gestures – at least outside a communicative context. Beat gestures are a quite complex visual stimulus, whereas the moving dots are a simple visual stimulus. The observed dissociation between visual beats and moving dots may be not so surprising when considering the communicative point of view. The beat gestures are most likely interpreted as a communicatively intended signal (Grice, 1975; the speaker wants me to pay particular attention to the phrase that accompanies the beat), whereas the moving points lack this communicative intention. We suggest that it is this difference in communicative intention that determines whether visual emphasis increases the plausibility of OSV structures (in the case of beat gestures) or not (in the case of moving dots).

Our experimental finding that a beat gesture influences the syntactic aspect of language makes an important and novel contribution to recent theorizing about the relationship between gesture and speech in comprehension. There already is convincing evidence that gesture and speech are tightly related in language production. We gesture when we speak, not when we listen (Levelt et al., 1985; McNeill, 1992); producing a gesture changes the pronunciation of the accompanying speech (Gentilucci et al., 2006; Krahmer and Swerts, 2007) and the syntactic properties of a language shape the form of their speakers' gestures (Kita and Özyürek, 2003). Thus, there is a bidirectional influence between gesture and speech in production. McNeill (1992) has argued that this interaction is so fundamental that gesture and speech together constitute language. On the basis of such production data, Kelly et al. (2010b) have recently put forward the idea that gesture and speech show a similar obligatory coupling during comprehension. In their integrated-systems hypothesis, they state that "gesture and speech mutually and obligatorily interact with one another to enhance language comprehension; that is, gesture influences the processing of speech, speech influences the processing of gesture, and this integration is mandatory." However, so far the integrated-systems hypothesis only applied to the semantic–conceptual level (Xu et al., 2009). The additional information present in gesture has been shown to facilitate language comprehension (Holle and Gunter, 2007; Holle et al., 2010; Wu and Coulson, 2010) and a semantically incongruent gesture–speech pairing interferes with gesture (Kelly et al., 2010b) as well as with language processing (Özyürek et al., 2007; Kelly et al., 2010a,b). Here, we show for the first time that an influence of gesture on speech also extends to the syntactic level. The present data reveal that the brain inevitably takes gesture into account when deciding which syntactic structure is the most plausible one for a given sentential input. Whether this relationship is also bidirectional (as has been shown for semantics) remains to be seen. Our data demonstrate an influence of gesture onto syntax. Future experiments will have to test whether there is also on influence of syntax onto gesture.

On a more general level, our results strongly support context-sensitive models of language processing. While previous research has shown that preceding context (Kaiser and Trueswell, 2004) as well as simultaneously presented visual scenes (Knoeferle et al., 2008) can influence which syntactic structure is assigned to a string of words, the present study extends these findings in several ways. First, unlike a sentence context or a visual scene, beat gestures do not operate on a semantic level. Instead, these hand movements can emphasize a certain phrase irrespective of their concrete form. It is just important at what time the movement occurred, not what it looked like. Second, unlike visual scenes, which are not normally part of face-to-face communication, gestures are always part of the communicative exchange, and may therefore serve as a very natural and powerful cue to shape the interpretation of spoken utterances.

In conclusion, we presented EEG evidence that simple beat gestures can enhance our understanding of syntactically more complex language. This finding has important consequences for our understanding of gesture and speech as integrated systems, and also has implications for everyday life. Most effective communication not only involves the mouth, but also the hands.

ACKNOWLEDGMENTS

This study was supported by the Max Planck Society. Henning Holle was supported by a grant from the Volkswagenstiftung (Az. II/82 175) during the preparation of the stimuli and subsequently by a grant from the ESRC awarded to Jamie Ward (Grant No RES-062-23-1150) during the preparation of the manuscript. We are grateful to Maria Bergau and Stefan Girgsdies for their help during the creation of the stimuli and Andrea Gast-Sandmann for assembling the figures.

SUPPLEMENTARY MATERIAL

The Movies S1–S15 for this article can be found online at http://www.frontiersin.org/Language_Sciences/10.3389/fpsyg.2012.00074/abstract

REFERENCES

Alibali, M. W., Heath, D. C., and Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: some gestures are meant to be seen. *J. Mem. Lang.* 44, 169–188.

Bader, M., and Meng, M. (1999). Subject-object ambiguities in German embedded clauses: an across-the-board comparison. *J. Psycholinguist. Res.* 28, 121–143.

Bavelas, J., Gerwing, J., Sutton, C., and Prevost, D. (2008). Gesturing on the telephone: independent effects of dialogue and visibility. *J. Mem. Lang.* 58, 495–520.

Boersma, P., and Weenink, D. (2010). *Praat: Doing Phonetics by Computer*, Version 5.1.15 [Computer program]. Available at: http://www.praat.org/

Crocker, M. W., Knoeferle, P., and Mayberry, M. R. (2010). Situated sentence processing: the coordinated interplay account and a neurobehavioral model. *Brain Lang.* 112, 189–201.

Efron, D. (1941/1972). *Gesture, Race and Culture*. The Hague: Mouton.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends Cogn. Sci. (Regul. Ed.)* 6, 78–84.

Friederici, A. D., Pfeifer, E., and Hahne, A. (1993). Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. *Brain Res. Cogn. Brain Res.* 1, 183–192.

Gentilucci, M., Bernardis, P., Crisi, G., and Dalla Volta, R. (2006). Repetitive transcranial magnetic stimulation of Broca's area affects verbal responses to gesture observation. *J. Cogn. Neurosci.* 18, 1059–1074.

Grice, H. P. (1975). "Logic and conversation," in *Speech Acts: Syntax and Semantics*, eds P. Cole and J. L. Morgan (New York: Academic Press), 41–58.

Hagoort, P., Brown, C., and Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Lang. Cogn. Process.* 8, 439–483.

Haupt, F. S., Schlesewsky, M., Roehm, D., Friederici, A. D., and Bornkessel-Schlesewsky, I. (2008). The status of subject-object reanalyses in the language comprehension architecture. *J. Mem. Lang.* 59, 54–96.

Holle, H., and Gunter, T. C. (2007). The role of iconic gestures in speech disambiguation: ERP evidence. *J. Cogn. Neurosci.* 19, 1175–1192.

Holle, H., Gunter, T. C., Ruschemeyer, S. A., Hennenlotter, A., and Iacoboni, M. (2008). Neural correlates of the processing of co-speech gestures. *Neuroimage* 39, 2010–2024.

Holle, H., Obleser, J., Rueschemeyer, S. A., and Gunter, T. C. (2010). Integration of iconic gestures and speech in left superior temporal areas boosts speech comprehension under adverse listening conditions. *Neuroimage* 49, 875–884.

Hubbard, A. L., Wilson, S. M., Callan, D. E., and Dapretto, M. (2009). Giving speech a hand: gesture modulates activity in auditory cortex during speech perception. *Hum. Brain Mapp.* 30, 1028–1037.

Kaiser, E., and Trueswell, J. C. (2004). The role of discourse context in the processing of a flexible word-order language. *Cognition* 94, 113–147.

Kelly, S. D., Creigh, P., and Bartolotti, J. (2010a). Integrating speech and iconic gestures in a Stroop-like task: evidence for automatic processing. *J. Cogn. Neurosci.* 22, 683–694.

Kelly, S. D., Özyürek, A., and Maris, E. (2010b). Two sides of the same coin: speech and gesture mutually interact to enhance comprehension. *Psychol. Sci.* 21, 260–267.

Kempen, G., and Harbusch, K. (2003). An artificial opposition between grammaticality and frequency: comment on Bornkessel, Schlesewsky, and Friederici. (2002). *Cognition* 90, 205–210; discussion 211–203, 215.

Kita, S., and Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *J. Mem. Lang.* 48, 16–32.

Knoeferle, P., Habets, B., Crocker, M. W., and Munte, T. F. (2008). Visual scenes trigger immediate syntactic reanalysis: evidence from ERPs during situated spoken comprehension. *Cereb. Cortex* 18, 789–795.

Krahmer, E., and Swerts, M. (2007). The effects of visual beats on prosodic prominence: acoustic analyses, auditory perception and visual perception. *J. Mem. Lang.* 57, 396–414.

Krauss, R. M. (1998). Why do we gesture when we speak. *Curr. Dir. Psychol. Sci.* 7, 54–60.

Krauss, R. M., Dushay, R. A., Chen, Y. S., and Rauscher, F. (1995). The communicative value of conversational hand gestures. *J. Exp. Soc. Psychol.* 31, 533–552.

Krauss, R. M., Morrel-Samuels, P., and Colasante, C. (1991). Do conversational hand gestures communicate? *J. Pers. Soc. Psychol.* 61, 743–754.

Levelt, W. J. M., Richardson, G., and la Heij, W. (1985). Pointing and voicing in deictic expressions. *J. Mem. Lang.* 24, 133–164.

McNeill, D. (1992). *Hand and Mind - What Gestures Reveal about Thought*. Chicago: The University of Chicago Press.

Obermeier, C., Holle, H., and Gunter, T. C. (2011). What iconic gesture fragments reveal about gesture-speech integration: when synchrony is lost, memory can help. *J. Cogn. Neurosci.* 23, 1648–1663.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.

Osterhout, L., and Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *J. Mem. Lang.* 31, 785–806.

Özyürek, A., Willems, R. M., Kita, S., and Hagoort, P. (2007). Online integration of semantic information from speech and gesture: insights from event-related brain potentials. *J. Cogn. Neurosci.* 19, 605–616.

Schriefers, H., Friederici, A. D., and Kuhn, K. (1995). The processing of locally ambiguous relative clauses in German. *J. Mem. Lang.* 34, 499–520.

van de Meerendonk, N., Kolk, H. H., Vissers, C. T., and Chwilla, D. J. (2010). Monitoring in language perception: mild and strong conflicts elicit different ERP patterns. *J. Cogn. Neurosci.* 22, 67–82.

Willems, R. M., Özyürek, A., and Hagoort, P. (2007). When language meets action: the neural integration of gesture and speech. *Cereb. Cortex* 17, 2322–2333.

Willems, R. M., Özyürek, A., and Hagoort, P. (2009). Differential roles for left inferior frontal and superior temporal cortex in multimodal integration of action and language. *Neuroimage* 47, 1992–2004.

Wu, Y. C., and Coulson, S. (2010). Gestures modulate speech processing early in utterances. *Neuroreport* 21, 522–526.

Xu, J., Gannon, P. J., Emmorey, K., Smith, J. F., and Braun, A. R. (2009). Symbolic gestures and spoken language are processed by a common neural system. *Proc. Natl. Acad. Sci. U.S.A.* 106, 20664–20669.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

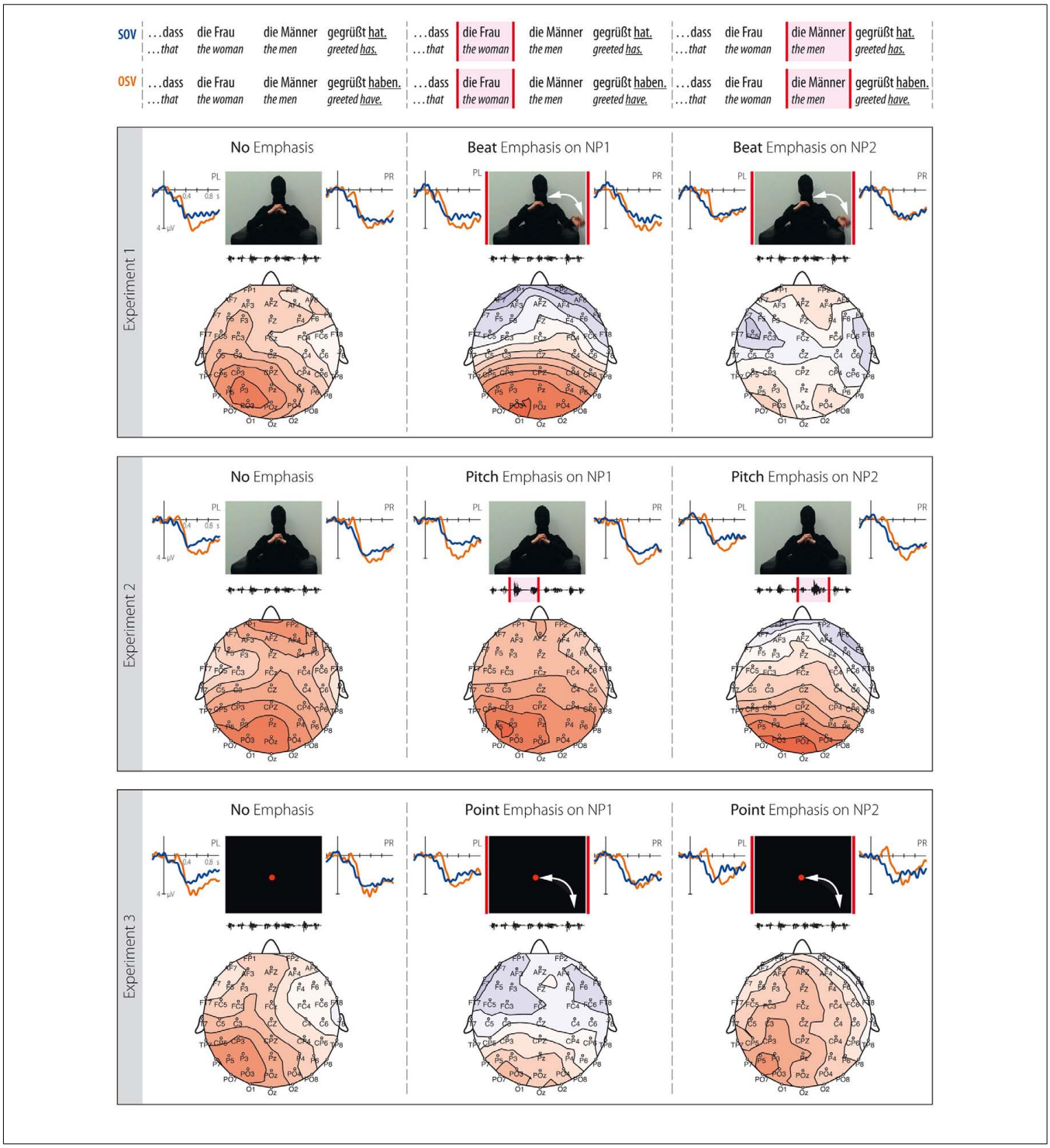
Received: 24 November 2011; paper pending published: 04 January 2012; accepted: 28 February 2012; published online: 19 March 2012.

Citation: Holle H, Obermeier C, Schmidt-Kassow M, Friederici AD, Ward J and Gunter TC (2012) Gesture facilitates the syntactic analysis of speech. Front. Psychology ***3****:74. doi: 10.3389/fpsyg.2012.00074*

This article was submitted to Frontiers in Language Sciences, a specialty of Frontiers in Psychology.

Copyright © 2012 Holle, Obermeier, Schmidt-Kassow, Friederici, Ward and Gunter. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.

APPENDIX


SOV
…dass die Frau die Männer gegrüßt hat.
…that the woman the men greeted has.
…dass die Frau die Männer gegrüßt hat.
…that the woman the men greeted has.
…dass die Frau die Männer gegrüßt hat.
…that the woman the men greeted has.
OSV
…dass die Frau die Männer gegrüßt haben.
…that the woman the men greeted have.
…dass die Frau die Männer gegrüßt haben.
…that the woman the men greeted have.
…dass die Frau die Männer gegrüßt haben.
…that the woman the men greeted have.
Experiment 1
No Emphasis
Beat Emphasis on NP1
Beat Emphasis on NP2
Experiment 2
No Emphasis
Pitch Emphasis on NP1
Pitch Emphasis on NP2
Experiment 3
No Emphasis
Point Emphasis on NP1
Point Emphasis on NP2
PL
PR

Movie no	Filename	Experiment	Structure	Emphasis	German sentence and literal English translation (emphasized part in bold)
1	Exp1_OI_PS_0.mpg	1	OSV	No emphasis	Simone denkt, dass die Kollegen die Denkerin gestört hat. *Simone thinks that the colleagues the thinker disturbed has.*
2	Exp1_OI_PS_1r.mpg	1	OSV	NP1 emphasis	Simone denkt, dass **die Kollegen** die Denkerin gestört hat. *Simone thinks that* ***the colleagues*** *the thinker disturbed has.*
3	Exp1_OI_PS_2r.mpg	1	OSV	NP2 emphasis	Simone denkt, dass die Kollegen **die Denkerin** gestört hat. *Simone thinks that the colleagues* ***the thinker*** *disturbed has.*
4	Exp1_SI_PS_0.mpg	1	SOV	No emphasis	Simone denkt, dass die Kollegen die Denkerin gestört haben. *Simone thinks that the colleagues the thinker disturbed have.*
5	Exp1_SI_PS_1r.mpg	1	SOV	NP1 emphasis	Simone denkt, dass **die Kollegen** die Denkerin gestört haben. *Simone thinks that* ***the colleagues*** *the thinker disturbed have.*
6	Exp1_SI_PS_2r.mpg	1	SOV	NP2 emphasis	Simone denkt, dass die Kollegen **die Denkerin** gestört haben. *Simone thinks that the colleagues* ***the thinker*** *disturbed have.*
7	Exp2_OI_PS_0.mpg	2	OSV	No emphasis	Simone denkt, dass die Kollegen die Denkerin gestört hat. *Simone thinks that the colleagues the thinker disturbed has.*
8	Exp2_OI_PS_1.mpg	2	OSV	NP1 emphasis	Simone denkt, dass **die Kollegen** die Denkerin gestört hat. *Simone thinks that* ***the colleagues*** *the thinker disturbed has.*
9	Exp2_OI_PS_2.mpg	2	OSV	NP2 emphasis	Simone denkt, dass die Kollegen **die Denkerin** gestört hat. *Simone thinks that the colleagues* ***the thinker*** *disturbed has.*
10	Exp2_SI_PS_0.mpg	2	SOV	No emphasis	Simone denkt, dass die Kollegen die Denkerin gestört haben. *Simone thinks that the colleagues the thinker disturbed have.*
11	Exp2_SI_PS_1.mpg	2	SOV	NP1 emphasis	Simone denkt, dass **die Kollegen** die Denkerin gestört haben. *Simone thinks that* ***the colleagues*** *the thinker disturbed have.*
12	Exp2_SI_PS_2.mpg	2	SOV	NP2 emphasis	Simone denkt, dass die Kollegen **die Denkerin** gestört haben. *Simone thinks that the colleagues* ***the thinker*** *disturbed have.*
13	Exp3_OI_PS_0.mpg	3	OSV	No emphasis	Simone denkt, dass die Kollegen die Denkerin gestört hat. *Simone thinks that the colleagues the thinker disturbed has.*
14	Exp3_OI_PS_1r.mpg	3	OSV	NP1 emphasis	Simone denkt, dass **die Kollegen** die Denkerin gestört hat. *Simone thinks that* ***the colleagues*** *the thinker disturbed has.*
15	Exp3_OI_PS_2r.mpg	3	OSV	NP2 emphasis	Simone denkt, dass die Kollegen **die Denkerin** gestört hat. *Simone thinks that the colleagues* ***the thinker*** *disturbed has.*
16	Exp3_SI_PS_0.mpg	3	SOV	No emphasis	Simone denkt, dass die Kollegen die Denkerin gestört haben. *Simone thinks that the colleagues the thinker disturbed have.*
17	Exp3_SI_PS_1r.mpg	3	SOV	NP1 emphasis	Simone denkt, dass **die Kollegen** die Denkerin gestört haben. *Simone thinks that* ***the colleagues*** *the thinker disturbed have.*
18	Exp3_SI_PS_2r.mpg	3	SOV	NP2 emphasis	Simone denkt, dass die Kollegen **die Denkerin** gestört haben. *Simone thinks that the colleagues* ***the thinker*** *disturbed have.*